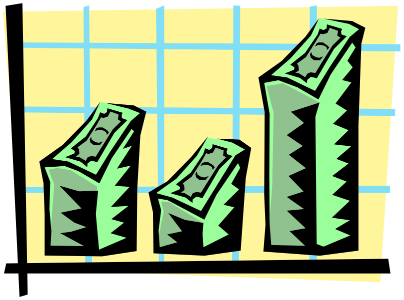
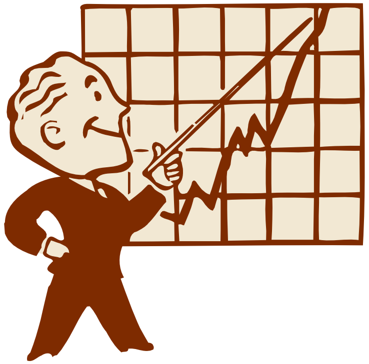




Introduction to Probability and Statistics

Slides 3 – Chapter 3

Ammar M. Sarhan,
asarhan@mathstat.dal.ca
Department of Mathematics and Statistics,
Dalhousie University
Fall Semester 2008

Chapter 3

Discrete Random Variables *and* Probability Distributions

Chapter Outlines

3.1 Random Variable

3.2 Probability Distribution for Discrete Random Variables

3.3 Expected Value of Discrete Random variables

3.4 Binomial Distribution

3.5 Hypergeometric & Negative Distributions

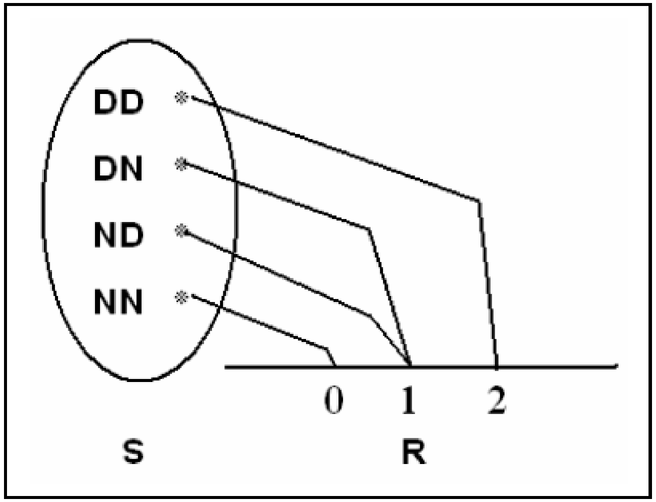
3.6 Poisson Distribution

In a statistical experiment, it is often very important to allocate numerical values to the outcomes.

Example:

- Experiment: testing two components (D=defective, N=non-defective)
- Sample space: $S=\{DD,DN,ND,NN\}$
- Let X = number of defective components.
- Assigned numerical values to the outcomes are:

Sample point (Outcome)	Assigned Numerical Value (x)
DD	2
DN	1
ND	1
NN	0



Notice that, the set of all possible values of the variable X is {0, 1, 2}.

Random Variable

For a given sample space S of some experiment, a *random variable is any* rule (function) that associates a real number with each outcome in S.

$$(\text{i.e., } X : S \rightarrow \mathbf{R}.)$$

Notation: " X " denotes the random variable .
" x " denotes a particular value of the random variable X.

Bernoulli Random Variable

Any random variable whose only possible values are *0* and *1* is called a *Bernoulli random variable*.

Types of Random Variables:

- A random variable X is called a discrete random variable if its set of possible values is countable, i.e., $x \in \{x_1, x_2, \ldots\}$
- A random variable X is called a continuous random variable if it can take values on a continuous scale, i.e., $x \in \{x: a < x < b;\ a, b \in R\}$

In most practical problems:

- A discrete random variable represents count data, such as the number of defectives in a sample of *k items*.
- A continuous random variable represents measured data, such as height.

3.2 Probability Distributions for Discrete Random Variables

Probability Distribution

The *probability distribution* or *probability mass function (pmf) of a* discrete rv X is defined for every number x by

$$p(x) = P\left(\text{all } s \in S : X(s) = x\right) = P\left(X = x\right)$$

The probability mass function, $p(x)$, of a discrete random variable X satisfies the following:

1) $p(x) = P\left(X = x\right) \geq 0.$

2) $\sum_{\text{all } x} p(x) = 1.$

Note:

$$\boxed{P(A) = \sum_{\text{all } x \in A} p(x) = \sum_{\text{all } x \in A} P(X = x).}$$

Parameter of a Probability Distribution

- Suppose that *p(x)* depends on a **quantity** that can be assigned any one of a number of possible values, each with different value determining a different probability distribution.
- Such a quantity is called a ***parameter of the distribution***.
- The collection of all distributions for all different parameters is called a ***family*** of distributions.

Example: For 0< *a* < 1,

$$p(x;a)=\begin{cases}1-a & if \;\; x=0\\ a & if \;\; x=1\\ 0 & otherwise\end{cases}$$

Each choice of *a* yields a different pmf.

Cumulative Distribution Function

The cumulative distribution function (cdf) *F(x) of a discrete rv variable X with pmf p(x) is defined by*

$$F(x) = P(X \le x) = \sum_{y:y\le x} p(y).$$

For any number *x*, *F(x) is the probability* that the observed value of *X will be at* most *x*.

Proposition

For any two numbers *a and b with a ≤ b*

$$\boxed{P(a \le X \le b) = F(b) - F(a-)}$$

"*a–*" *represents the largest possible X* value that is strictly less than *a*.

Note: For integers

$$P(a \le X \le b) = F(b) - F(a-1)$$

Probability Distribution for the Random Variable *X*

A probability distribution for a random variable *X*:

x	-3	-2	-1	0	1	4	6
P(X=x)	0.13	0.16	0.17	0.20	0.16	0.11	0.07

Find:

1)F(x)

2)$P(X > 0)$ $= 1 - P(X \le 0) = 1 - F(0) = 1 - 0.66 = 0.34$

3)$P(-2 \le X \le 1)$ $= P(-2 \le X \le 1) = F(1) – F(-2-) = F(1) – F(-3) = 0.69$

4)$P(-2 \le X < 1)$ $= P(-2 \le X < 1) = F(1-) – F(-2-) = F(0) – F(-3) = 0.53$

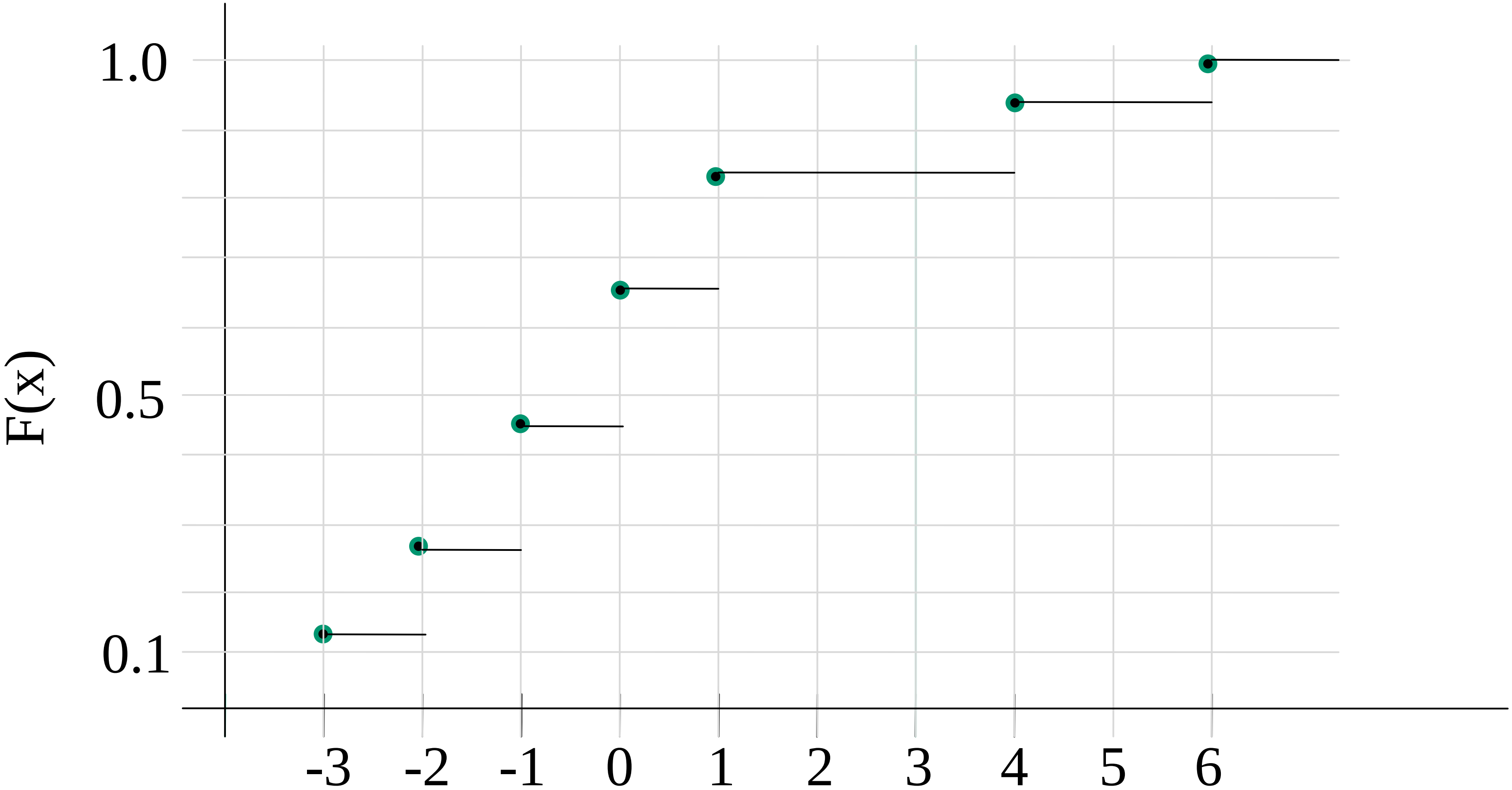
5)$P(-2 < X \le 1)$ $= P(-2 < X \le 1) = F(1) – F(-2) = F(1) – F(-2) = 0.53$

6)$P(-3 < X < 1)$ $= P(-3 < X < 1) = F(1-) – F(-2) = F(0) – F(-2) = 0.\ 37$

1) F(x)

x	-3	-2	-1	0	1	4	6
F(x)	0.13	0.29	0.46	0.66	0.82	0.93	1.00

x	-3	-2	-1	0	1	4	6
F(x)	0.13	0.29	0.46	0.66	0.82	0.93	1.00



x	-3	-2	-1	0	1	4	6
P(X=x)	0.13	0.16	0.17	0.20	0.16	0.11	0.07

pmf:

x	-3	-2	-1	0	1	4	6
P(X=x)	0.13	0.16	0.17	0.20	0.16	0.11	0.07

CDF:

x	-3	-2	-1	0	1	4	6
F(x)	0.13	0.29	0.46	0.66	0.82	0.93	1.00

The CDF can be written as

$$F(x) = \begin{cases} 0 & \text{if } x < -3 \\ 0.13 & \text{if } -3 \le x < -2 \\ 0.29 & \text{if } -2 \le x < -1 \\ 0.46 & \text{if } -1 \le x < 0 \\ 0.66 & \text{if } 0 \le x < 1 \\ 0.82 & \text{if } 1 \le x < 4 \\ 0.93 & \text{if } 4 \le x < 6 \\ 1.00 & \text{if } 6 \le x \end{cases}$$

Example.

Tossing a non-balance coin 2 times independently.

Sample space: *S={HH, HT, TH, TT}*.

Suppose P(H) = ½ P(T) $\Longrightarrow$ P(H) =1/3 and P(T)=2/3.

Let X= number of heads

Sample point	Value of X (x)	Probability
HH	2	P(HH) = P(H) P(H) = 1/3×1/3=1/9
HT	1	P(HT) =P(H) P(T) =1/3×2/3 = 2/9
TH	1	P(HT) = P(H) P(T) =1/3×2/3 = 2/9
TT	0	P(TT) = P(T) P(T) = 2/3×2/3 = 4/9

The possible values of X are: 0, 1, and 2.

X is a discrete random variable.

(X=x)	p(x) = P(X=x)
(X=0)={TT}	4/9
(X=1)={HT, TH}	2/9 + 2/9 =4/9
(X=2)={HH}	1/9

$\Rightarrow$

Probability distribution of the X

X	0	1	2	Total
P(X=x)	4/9	4/9	1/9	**1.0**

$P(X<1) = P(X=0)=4/9$

$= F(1-) = F(0) = 4/9$

$P(X\leq 1) = P(X=0) + P(X=1) = 4/9+4/9 = 8/9$

$= F(1) = 8/9$

$P(X\geq 0.5) = P(X=1) + P(X=2) = 4/9+1/9 = 5/9$

$= 1-F(0.5-) = 1-F(0) = 1-4/9 = 5/9$

$P(X>8) = P(\varphi) = 0$

$=1-F(8) = 1-1 = 0$

$P(X<10) = P(X=0) + P(X=1) + P(X=2) = P(S) = 1$

$= F(10) = 1$

$$\boxed{F(x)=\begin{cases}0 & \text{if } x<0\\ 4/9 & \text{if } 0\leq x<1\\ 8/9 & \text{if } 1\leq x<2\\ 1 & \text{if } 2\leq x\end{cases}}$$

F(x)

X	0	1	2
F(x)	4/9	8/9	9/9

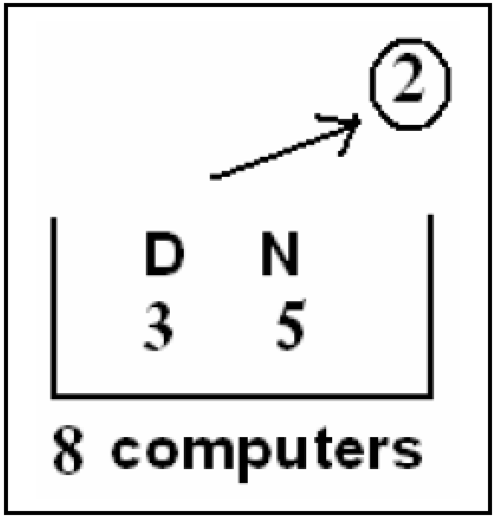
Example :

A shipment of 8 similar microcomputers to a retail outlet contains 3 that are defective and 5 are non-defective. If a school makes a random purchase of 2 of these computers, *find the probability distribution of the number of defectives* .

We need to find the probability distribution of the random variable: X = the number of defective computers purchased.

Experiment: selecting 2 computers at random out of 8

$$N(S) = \binom{8}{2} \text{ equally likely outcomes}$$



The possible values of X are: x = 0, 1, 2.

$$N(X=0)=\{0D \text{ and } 2N\} = \binom{3}{0}\times\binom{5}{2}$$

$$N(X=1)= \{1D \text{ and } 1N\} = \binom{3}{1}\times\binom{5}{1}$$

$$N(X=2)=\{2D \text{ and } 1N\} = \binom{3}{2}\times\binom{5}{0}$$

$$P(X=0)= \frac{\binom{3}{0}\times\binom{5}{2}}{\binom{8}{2}}=\frac{10}{28}$$

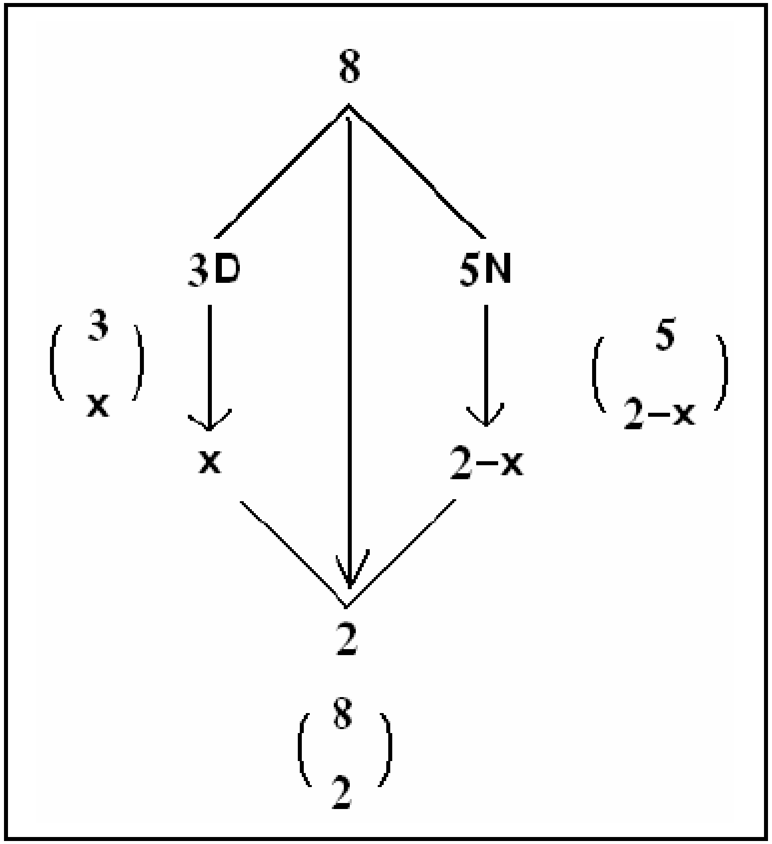
$$P(X=1)= \frac{\binom{3}{1}\times\binom{5}{1}}{\binom{8}{2}}=\frac{15}{28}$$

$$P(X=2)= \frac{\binom{3}{2}\times\binom{5}{0}}{\binom{8}{2}}=\frac{3}{28}$$

The probability distribution of X

x	0	1	2	Total
p(x) = P(X=x)	$\frac{10}{28}$	$\frac{15}{28}$	$\frac{3}{28}$	1.00

$$p(x) = \begin{cases} \dfrac{\binom{3}{x}\times\binom{5}{2-x}}{\binom{8}{2}} & , x = 0,1,2. \\ 0, & \text{otherwise.} \end{cases}$$



Hypergeometric Distribution

3.3 Expected Values of Discrete Random Variables

The Expected Value of *X*

Let X be a discrete rv with set of possible values D and *pmf* $p(x)$. The expected value or mean value of X, denoted $E(X)$ or μ_X or μ, is

$$E(X) = \mu_X = \sum_{x \in D} x \cdot p(x).$$

Example: Use the data below to find out the expected number of the number of credit cards that a student will possess.

X = # credit cards

x	0	1	2	3	4	5	6
p(x)	0.08	0.28	0.38	0.16	0.06	0.03	0.01

$$E(X) = x_1 p_1 + x_2 p_2 + \dots + x_1 p_1$$

$$= 0(0.08) + 1(0.28) + 2(0.38) + 3(0.16) + 4(0.06) + 5(0.03) + 6(0.01)$$

$= 1.97 \approx 2$ credit cards

The Expected Value of a Function

If the rv X has the set of possible values D and pmf $p(x)$, then the expected value of any function $h(x)$, denoted E[h(X)] or ($\mu_{h(X)}$) is

$$E[h(X)] = \mu_{h(X)} = \sum_{x \in D} h(x) \cdot p(x).$$

Rules of the Expected Value

$$\boxed{\mathrm{E}[aX + b] = a \cdot \mathrm{E}(X) + b}$$

1) For any constant a, $\mathrm{E}[aX] = a \cdot \mathrm{E}(X)$

2) For any constant b, $\mathrm{E}[X + b] = \mathrm{E}(X) + b$

The Variance and Standard Deviation

Let X have pmf $p(x)$, and expected value μ. Then the **variance** of X, denoted $V(X)$ (or σ_X^2 *or* σ^2), is

$$\boxed{V(X) = E[(X-\mu)^2] = \sum_{x \in D} (x-\mu)^2 \cdot p(x).}$$

The ***standard deviation*** (SD) of X is

$$\boxed{\sigma_X = \sqrt{\sigma_X^2}}$$

Example: The quiz scores for a particular student are given below:

22, 25, 20, 18, 12, 20, 24, 20, 20, 25, 24, 25, 18

Compute the variance and standard deviation.

Solution: Let X be the quiz score.

$p(x) \longrightarrow$

Value (X)	12	18	20	22	24	25
Frequency	1	2	4	1	2	3
Probability	0.08	0.15	0.31	0.08	0.15	0.23

$$\mu = \sum_{x \in D} x \cdot p(x) = 21$$

$$V(X) = (x_1 - \mu)^2 \, p_1 + (x_1 - \mu)^2 \, p_1 + \ldots + (x_n - \mu)^2 \, p_n$$

$$= (12 - 21)^2 \, 0.08 + (18 - 21)^2 \, 0.15 + (20 - 21)^2 \, 0.31 +$$

$$(22 - 21)^2 0.08 + (24 - 21)^2 \, 0.15 + (25 - 21)^2 \, 0.23 = 13.25$$

$$\sigma = \sqrt{V(X)} = \sqrt{13.25} \approx 3.64$$

21

Shortcut Formula for Variance

$$V(X)=\sigma^2=\left[\sum_D x^2\cdot p(x)\right]-\mu^2$$

$$=E\left(X^2\right)-\left[E(X)\right]^2$$

Rules of Variance

$$V(aX+b)=\sigma^2_{aX+b}=a^2\cdot\sigma^2_X$$

and $\sigma_{aX+b}=|a|\cdot\sigma_X$

This leads to the following:

1. $\sigma^2_{aX}=a^2\cdot\sigma^2_X,\ \sigma_{aX}=|a|\cdot\sigma_X$
2. $\sigma^2_{X+b}=\sigma^2_X$

22

3.4 The Binomial Probability Distribution

Binomial Experiment

An experiment for which the following four conditions are satisfied is called a ***binomial experiment***.

1. The experiment consists of a sequence of n trials, where n is fixed in advance of the experiment.
2. The trials are identical, and each trial can result in one of the same two possible outcomes, which are denoted by success (*S)* or failure *(F)*.
3. The trials are independent
4. The probability of success is constant from trial to trial: denoted by p.

Binomial Random Variable

Given a binomial experiment consisting of *n trials*, ***the binomial random variable*** *X* associated with this experiment is defined as
X = the number of S's among n trials.

Notation for the pmf of a Binomial rv

Because the pmf of a binomial rv *X* depends on the two parameters *n* and *p*, we denote the pmf by *b(x;n,p).*

Computation of a Binomial pmf

$$b(x;n,p) = \begin{cases} \binom{n}{p} p^{x}(1-p)^{n-x} & x = 0,1,\cdots,n \\ 0 & \text{otherwise} \end{cases}$$

The expected value and Variance of a Binomial distribution

If $X \sim b(x;n,p)$ (1) $E(X) = n\,p$ (2) $V(X) = n\,p\,(1-p)$

Example: If the probability of a student successfully passing this course (C or better) is 0.82, find the probability that given 8 students

a. all 8 pass.

$$P(X = 8) = b(8;8,0.82) = \binom{8}{8}(0.82)^8(1-0.82)^{8-8} \approx 0.2044$$

b. none pass.

$$P(X = 0) = b(0;8,0.82) = \binom{8}{0}(0.82)^0(1-0.82)^{8-0} \approx 0.0000011$$

c. at least 6 pass.

$$P(X \geq 6) = b(6;8,0.82) + b(7;8,0.82) + b(8;8,0.82)$$

$$= \binom{8}{6}(0.82)^6(1-0.82)^{8-6} + \binom{8}{7}(0.82)^7(1-0.82)^{8-7}$$

$$\binom{8}{8}(0.82)^8(1-0.82)^{8-8} \approx 0.8392$$

d. The expected number of students passed the course.

$$E(X) = n\,p = 8\,(0.82) = 6.56 \approx 7 \text{ students}$$

e. The variance.

$$\begin{aligned} V(X) &= n\,p\,(1-p) \\ &= 8\,(0.82)\,(1-0.82) = 8\,(0.82)\,(0.18) \\ &= 1.1808 \end{aligned}$$

26

The Hypergeometric Distribution

The three assumptions that lead to a *hypergeometric distribution*:

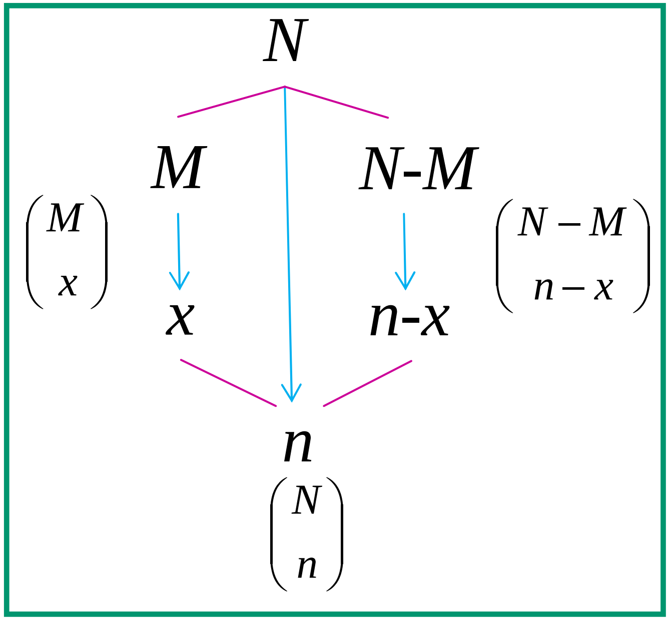
1. The population or set to be sampled consists of N individuals, objects, or elements (a finite population).
2. Each individual can be characterized as a success (S) or failure (F), and there are M successes in the population.
3. A sample of n individuals is selected without replacement in such a way that each subset of size n is equally likely to be chosen.

If X is the number of S's in a completely random sample of size n drawn from a population consisting of M S's and $(N - M)$ F's, then the probability distribution of X will be called the hypergeometric distribution. [$h\ (n,M,N)$]

Computation of a Hypergeometric distribution

$$P(X=x)=h(x;n,M,N)=\frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}},$$

$$\max(0,n-N+M)\le x\le \min(n,M)$$



The expected value and Variance of a Hypergeometric distribution

If $X\sim h\,(n,M,N)$

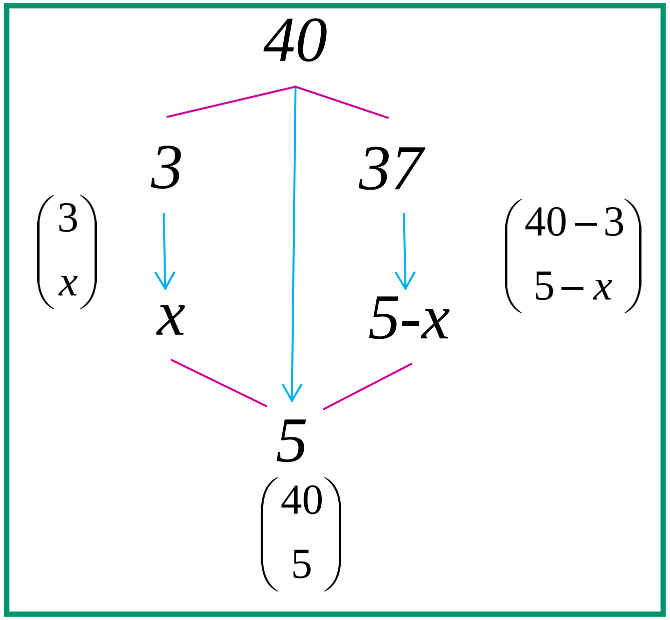
(1) $E(X)=n\dfrac{M}{N}$

(2) $V(X)=\left(\dfrac{N-n}{N-1}\right)\cdot n\cdot\dfrac{M}{N}\cdot\left(1-\dfrac{M}{N}\right)$

Example:

Lots of 40 components each are called acceptable if they contain no more than 3 defectives. The procedure for sampling the lot is to select 5 components at random (without replacement) and to reject the lot if a defective is found. What is the probability that exactly one defective is found in the sample if there are 3 defectives in the entire lot.

Solution:



Let X= number of defectives in the sample.

X ~ h(n, M, N) ≡ h(5, 3, 40)

$N = 40$, $M = 3$, $n = 5$.

$$P(X = x) = \frac{\binom{3}{x}\binom{37}{5-x}}{\binom{40}{5}}, \quad x = 0, 1, 2, 3$$

$$\max(0, 5-37) = 0 \le x \le \min(5, 3) = 3$$

29

The probability that exactly one defective is found in the sample is

$$P(X=1) = \frac{\binom{3}{1}\times\binom{37}{5-1}}{\binom{40}{5}} = \frac{\binom{3}{1}\times\binom{37}{4}}{\binom{40}{5}} = 0.3011$$

Example:

In the previous example, find the expected value and the variance of the number of defectives in the sample.

Solution:

$$(1)\ E(X) = n\frac{M}{N} = 5\frac{3}{40} = \frac{3}{8} = 0.375$$

$$(2)\ V(X) = \left(\frac{N-n}{N-1}\right)\cdot n\cdot\frac{M}{N}\cdot\left(1-\frac{M}{N}\right)$$

$$= \left(\frac{40-5}{40-1}\right)\cdot 5\cdot\frac{3}{40}\cdot\left(1-\frac{3}{40}\right) = 0.311298$$

The Negative Binomial Distribution

The *negative binomial rv* and distribution are based on an experiment satisfying the following four conditions:

1. The experiment consists of a sequence of independent trials.
2. Each trial can result in a success (S) or a failure (F).
3. The probability of success is constant from trial to trial, so $P(S \text{ on trial } i) = p$ for $i = 1, 2, 3, \ldots$
4. The experiment continues until a total of r successes have been observed, where r is a *specified positive integer*.

If X is the number of failures that precede the r-th success, then the probability distribution of X will be called the negative binomial distribution. [$nb\,(r, p)$]

The event ($X = x$) is equivalent to {r-1 S's in the first (x+r-1) trials and an S in the (x+r)th}.

Computation of a Negative Binomial pmf

$$nb(x; r, p) = \binom{x+r-1}{r-1} p^r (1-p)^x, \quad x = 0,1,2,\cdots$$

The expected value and Variance of a Negative Binomial distribution

If $X \sim nb(r, p)$ ⇨ (1) $E(X) = r\,(1-p)/p$ (2) $V(X) = r\,(1-p)/p^2$

Geometric Distribution:

When r = 1 in nb(r, p) then the nb distribution reduces to geometric distribution.

Example:

Suppose that p=P(male birth) = 0.5. A parent wishes to have exactly two female children in their family. They will have children until this condition is fulfilled.

1. *What is the probability that the family has x male children?*

   Let X be the number of Ms precede the *2nd F.* $\boxed{X \sim nb(2, 0.5)}$

   The prob. that the family has x male children = $P(X=x)$

   $\boxed{= nb(x;\ 2,0.5),\ x=0,1,2,\ldots}$

2. *What is the probability that the family has 4 children?*

   $\boxed{= P(X = 2) = nb(2;\ 2,\ 0.5) = 0.1875}$

3. *What is the probability that the family has at most 4 children?*

   $= P(X \le 2) = P(X = 0) + P(X = 1) + P(X = 2)$

   $= nb(0;\ 2,\ 0.5) + nb(1;\ 2,\ 0.5) + nb(2;\ 2,\ 0.5)$

   $= 0.6875$

4. *How many male children would you expect this family to have? How many children would you expect this family to have?*

$E(X) = r\,(1 - p)/\,p = 2\,(1\text{-}0.5)\,/\,0.5 = 2$

$2 + E(X) = 2 + 2 = 4$

3.6 The Poisson Probability Distribution

A random variable X *is said to have* a ***Poisson distribution*** *with* parameter λ (λ > 0), if the pmf of X is

$$p(x;\lambda) = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x = 0,1,2,\cdots$$

The Poisson Distribution as a Limit

Suppose that in the binomial pmf $b(x;n, p)$, we let $n \to \infty$ and $p \to 0$ in such a way that np approaches a value λ (λ > 0). Then

$$\boldsymbol{b(x;\, n,\, p) \to p(x;\, \lambda)}$$

The expected value and Variance of a Poisson distribution

If X has a Poisson distribution with parameter λ , then

$$\boxed{E(X) = V(X) = \lambda}$$

Poisson Process

Assumptions:

1. There exists a parameter $\alpha > 0$ such that for any short time interval of length Δt , the probability that exactly one event is received is

$$\alpha \Delta t + o(\Delta t)$$

2. The probability of more than one event during Δt is $o(\Delta t)$.
3. The number of events during the time interval Δt is independent of the number that occurred prior to this time interval.

Poisson Process

$P_k(t) = e^{-\alpha t} (\alpha t)^k / k!$, so that the number of pulses (events) during a time interval of length t is a Poisson rv with parameter αt. The expected number of pulses (events) during any such time interval is αt, so the expected number during a unit time interval is α.

Example:

Suppose that the number of typing errors per page has a Poisson distribution with average 6 typing errors.

(1) What is the probability that in a given page:

(i) The number of typing errors will be 7?

(ii) The number of typing errors will be at least 2?

Let X = number of typing errors per page. $\Rightarrow$ X ~ Poisson (6)

$$\Rightarrow \quad P(X=x)=p(x;6)=\frac{6^{x}e^{-6}}{x!},\quad x=0,1,2,\cdots$$

(i) $$P(X=7)=p(7;6)=\frac{6^{7}e^{-6}}{7!}=0.13768.$$

(ii) $$P(X\geq 2)=1-P(X<2)=1-\left[P(X=0)+P(X=1)\right]$$

$$=1-p(0;6)-p(1;6)=0.982650$$

Dr. Ammar M. Sarhan 37

(2) What is the probability that in 2 pages there will be 10 typing errors?

Let X = number of typing errors in 2 pages. ⟹ X ~ Poisson (12)

$$P(X = 10) = p(10;12) = \frac{12^{10} e^{-12}}{10!} = 0.1084.$$

(3) What is the probability that in a half page there will be no typing errors?

Let X = number of typing errors in a half page. ⟹ X ~ Poisson (3)

$$P(X = 0) = p(0;3) = \frac{3^{0} e^{-3}}{0!} = 0.0497871.$$